Pergamon

Insect Biochemistry and Molecular Biology 32 (2002) 951–960

Insect Biochemistry and Molecular Biology

www.elsevier.com/locate/ibmb

Evolution and phylogenetic information content of the ribosomal DNA repeat unit in the Blattodea (Insecta)

Dmitrv Mukha [a] *, Brian M. Wiegmann [b], Coby Schal [b]

[a] *Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin St. 3, 117809, Moscow, Russia*
[b] *Department of Entomology, Box 7613, North Carolina State University, Raleigh, NC 27695, USA*

Received 5 December 2000; received in revised form 25 October 2001; accepted 26 October 2001

Abstract

The organization, structure, and nucleotide variability of the ribosomal repeat unit was compared among families, genera, and species of cockroaches (Insecta: Blattodea). Sequence comparisons and molecular phylogenetic analyses were used to describe rDNA repeat unit variation at differing taxonomic levels. A ~1200 bp fragment of the 28S rDNA sequence was assessed for its potential utility in reconstructing higher-level phylogenetic relationships in cockroaches. Parsimony and maximum likelihood analyses of these data strongly support the expected pattern of relationships among cockroach groups. The examined 5′ end of the 28S rDNA is shown to be an informative marker for larger studies of cockroach phylogeny. Comparative analysis of the nucleotide sequences of the rDNA internal transcribed spacers (ITS1 and ITS2) among closely related species of *Blattella* and *Periplaneta* reveals that ITS sequences can vary widely in primary sequence, length, and folding pattern. Secondary structure estimates for the ITS region of *Blattella* species indicate that variation in this spacer region can also influence the folding pattern of the 5.8S subunit. These results support the idea that ITS sequences play an important role in the stability and function of the rRNA cluster. © 2002 Published by Elsevier Science Ltd.

Keywords: Phylogenetic analysis; 28S ribosomal DNA; Ribosomal internal transcribed spacer 1 and 2; Blattodea

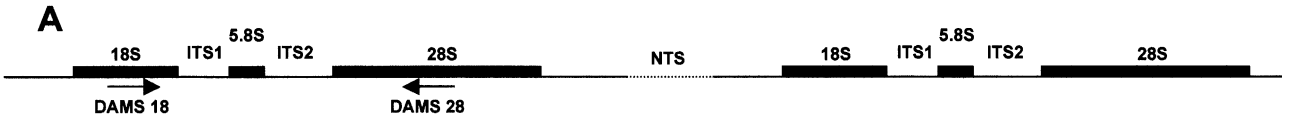
1. Introduction

Ribosomal DNA (rDNA) has long been considered a useful marker for comparative evolutionary and phylogenetic studies (Hillis and Davis, 1986; Mindell and Honeycutt, 1990; Wesson et al., 1993; Schlotterer et al., 1994; Vogler and DeSalle, 1994; Tang et al., 1996). This utility results largely from the differential conservation and variability of different genes and regions that make up the rDNA repeat unit (Hillis and Dixon, 1991; Hamby and Zimmer, 1992). The basic organization of rDNA is conserved throughout eukaryotes. Eukaryotic ribosomal RNA genes are arranged in tandemly repeated clusters, with each cluster containing the genes for 18S-, 5.8S-, and 28S-like ribosomal RNAs. The genes are separated by several spacers, namely, the NTS (nontranscribed spacer), and ITS1 and ITS2 (internal transcribed spacers). The NTS separates neighboring repeat units; ITS1 is located between the 18S- and 5.8S-like coding regions; ITS2 lies between the 5.8S- and 28S-like genes (Fig. 1A) (Gerbi, 1985). Ribosomal repeats are usually localized to one or a few chromosomes and form part of the nucleolar organizers (NO). Because the coding regions and spacers differ widely in their rate of evolution, they can reveal phylogenetic relationships ranging from the level of major phyla of living organisms to the population level (Hillis and Dixon, 1991; Wesson et al., 1992; Kuperus and Chapco, 1994; Honda et al., 1998; Muccio et al., 2000; Wiegmann et al., 2000).

Additional information about relationships and rDNA evolution can be obtained by examining the secondary structure of the ribosomal RNA. In the rDNA coding regions (18S-, 5.8S-, and 28S-like), sequence conservation is reflected in the high similarity of secondary structures in a variety of distantly related organisms (Gerbi et al., 1982; Raué et al., 1988; Wesson et al., 1992). Sequence tracts that are highly variable between

*Corresponding author. Address for correspondence: Department of Entomology, Box 7613, North Carolina State University, Raleigh, NC 27695, USA.

0965-1748/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0965-1748(01)00164-3



B

```
CTACTAGATGGTTCGATTAGTCTTTCGCCCCTATACCCAGCTCTGACGATCGATTTGCACGTCAGACTCGCTGCG   DAMS 28
GACCTCCATCAGGGTTTCCCCTGACTTCGTCCTGACCAGGCATAGTTCACCATCTTTCGGGTCCCAACGTGTACG   coc 4
CTCTGGGTGCGCCTGCGGCCGAGGCCGAGGCGCCCCGGGAGTGCGGGCCCGGACACTCGCCGCGGAGGGCGAGGC
CAGACGGCCATCCTCCCAACAAACATCGGCAACGCCACGAGTGACGCGCCTTTGCCATGCTTTCACTGCGCCTTT
GGGTTTAGACAGCCCAATGACTCGCGCACATGTTAGACTCCTTGGTCCGTGTTTCAAGACGGGTCGGGAAAGAGA   coc 1
CCGAAGCTTTAAGCGCCGCCGACGGGAGACTGCTACCGTCCGAGGAATCCGAGGACCAACAGCCGACGGGGGCCG
GGGGCGTGGCGTTCGCCACGGACCACCCAGACACAACCCGAAGAGCTTGCCCTCGGTCCGGACGCTCACTCGCAG
TTAAGGAGCGGGCCTTCTCCGCCATGCTATACCGTCGGGCTCCGGCCAGGCATGGAACGGGTTCCCGGCGACCTT
TGGCCCGAGCCAAACCGGCTTCGGAGGTCGACTAAGGACAATTCCCAGAGCCCTTACAGACCGGTGCCCAACGAG
TCGCGACGTCCTACTCGGGGAAGAAGTGCACCTCGCCAGCACCAGAGACTTTGTGCTTCCGGGCCAAGGGTGACT
CCGGACCGACACGGAGCGCGCCTTGCGGCGCGACCGCTGGCCGGGGCCATCCCCAAGGCAACTTCGGGCGCCCTG
CGTGGACAAGAATGAGTCTCCCCCTTTCGCCTATTCGGATCTGTCCCGTTTACCCTTGAACGGTTTCACGTACTG
TTGAACTCTCTCTTCAAAGTTCTTTTCAACTTTCCCTCACGGTACTTGTTCGCTATCGGTCTCGTGGTCATATTT   coc 2
AGCCTTAGATGGAGTTTACCACCCACTTAGTGCTCCACTCTCAAGGAACACGACTCTGAGGAGAGCTCCTCTCGG
GCACGGTGCCGGTCACTACGGGCCTGGCACCCTCTACGGGATGTAGCCCCGTTCAAGGGAGACTTGAACTCGGCA   coc 3
CCGACGCCCGAGAATGCGGAACCTCCCACACACGACATCTCCCTGCGCGTTGCCACGCGGGATTCCGTGCTGGGC
TCTTCCCTGTTCGCTCGCCGCTACTTGGGGAATCCTTGTTAGTTTCTTTTCCTCCGCCTATTAATATGCTTAAAT   coc 6
TCAGCGGGTAGTCTCGCCTGATCTGGGGTCGGAAACAATTAAGACTGCCAAGAGCCTTGCCAGAAGCTTTTGCGT
CGAGCAGACGAACACCGGGGGTCGGTCACAAGGCTTTTTCTGTTACAAAGCAGCCAGAACCGCGCGCGCAGCGGA
TGTTCTTTGTCCGATCGACCCCTATCGGCTTTGCAGCCTCACGCGGGACCGACCATATAGCGGACGAGACAGAAC
CTTTTCCACACACGCATTGCCCCGTTCAGCATAAGCTGCCGACAGCTTCACGAGAGACTTAAAGAGCGACGGGCT
CGCAAAAGAAAGAAAAACAAATTTTTTTCCTTTTTCGCAAGGCACCGGTGCCCCCAAAGCGGACAAGCGACGGGC
CTCTCATTCCCACAGCAAGCGACACACAGTGTGTCGAGTCAGTGGAGAGAAAACGCCGACGATGCCACAGGAATT
GTTTGCAAGCCGACCCTCAGCCAGGCGtggtccgggattggaaatccaaggaccgcaatgtgcgttcgaaatgtc
gatgttcatgtgtcctgcagttcacatgtcgacgcgcaatttgctgcgttcttcatcgaccCACGAGCCGAGTGA   coc 5
TCCACCGttcagggttgtaATGATAGTTCATATAAATCCAAAACGACGCGCAAGGCGCCGAGAAGGAAATATAGA
AAAAATTTGCCAGACAAGGTATAAGTTTCGCCCCTCCGCGAGGAGGGGCTCCTTCGCATCCGCCACCAAGAGCCC
TAACCCCTCACACCACACAACGACTTCTCGAGCGATCGAGAGCGATGCGATGTACAGAGAGAGTAAGAGACAAGC
GAGACAGCGGACGAAAGGGGTCTTTGCCGAGGGCGAAGCCTAAAGCTGGCTCCTTTAAACCGCGCATTTTGCAGT
GCCCAAAGCATTCGACAGTAAAAAAAAAGGGCGAGTTAGGTCAGACGCAGCACGCGAAGCATGCAGGACAACCGC
CGCGCCACAAGGCGGTTACCAGACGCAGCACGTGAAGCATGCAGACCCTTACTACAAACCCTTTGGGCGAAGACT
GTTAGAGTCAGTCTTCTTTCAACGGCGCACACAAAAGCGCCGTACCCATAAAAAGGGCAACCTTCTCCGCCGAAA
CGGACGAAAGAGTTGCCTAAGTAAAAAACAATGTGCTCAAAAGCCCTTCCAGTAAAAACAGAATTCGCCCGCGCG
CGCGCGTCCACCAACGAACGACACACACGGGGAATCTCTCTCTCCTCCACAAAACTTCCATAGTGGCAACGATAA
ATTCCACGTTAATGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTTTACTTCCTCTAAATAATCA   coc 7
AGTTTGGGCATCTTCCCGGTGGAACGTGGGAGCAAGCCCCCGCGCCCGCCGGTCCGAAGACCTCACTAAATCATT
CAATCGGTAGTAGCGACGGGCGGTGTGTACAAACGGCAGGGGC                                    DAMS 18
```

Fig. 1. (A) Scheme of the eukaryotic tandemly repeated rDNA clusters; 18S, 5.8S, 28S — corresponding ribosomal RNA genes; ITS1 and ITS2 — corresponding internal transcribed spacers; NTS — nontranscribed spacer. (B) Nucleotide sequence of the *Blattella germanica* analyzed fragment and primer locations. Main structural elements of the analyzed fragment are indicated using various fonts: 5′ fragment of the 28S rDNA — capital bold; 3′ fragment of the 18S rDNA — capital bold italic; 5.8S rDNA — lower case bold; ITS1 and ITS2 — normal capital letters. The boundaries of the described structural elements were defined by comparison with corresponding sequences from GenBank. Sequences corresponding to the primers used for amplification and sequencing (indicated on the right) are underlined.

species may still retain certain, apparently functionally important, components of secondary structure. For example, estimated secondary structures of the ITS1 sequences are nearly identical in human, chimpanzee, and gorilla (Gonzales et al., 1990), and those from mouse and rat also show high similarity (Michot et al., 1983). Integrity of the rDNA secondary structure is maintained through compensatory nucleotide changes to preserve base pairing in stems, and through indels which change stem length but do not alter the overall folding pattern. Structural analysis of the ITS in yeast has shown that the spacer regions also play a role in the maturation of precursor rRNA molecules (Musters et al., 1990; Van der Sande et al., 1992; Schulenburg et al., 1999).

Despite the importance of rDNAs in phylogenetic and molecular evolutionary studies of insects, very little is known about the rDNA of cockroaches. Because of their pest status and the ease with which they are cultured in the laboratory, cockroaches have served as model organisms for studies of insect physiology, molecular genetics,

and chemical ecology, but relatively few of these studies have been explicitly comparative or evolutionary in their focus. Our studies of the nuclear ribosomal genes in cockroaches are aimed at understanding the dynamics of sequence and structural variation in these genes and assessing the utility of that variation for comparative evolutionary studies. In this paper, we examine the organization, structure, and nucleotide variability of the ribosomal repeat unit of cockroaches (Insecta: Blattodea). Our comparison focuses on the ITS1 and ITS2 spacer regions, as well as on portions of the rDNA subunit coding regions that immediately flank them. These regions are compared across a broad range of taxonomic divergences within cockroaches.

Divergences between extant genera of cockroaches may be as old as 75–100 my before present (Labandeira, 1994; Nalepa and Bandi, 1999), but some genus and species-level divergences could be much more recent (Nalepa and Bandi, 1999). Phylogenetic relationships among cockroach groups are the subject of current debate in the insect systematics literature (McKittrick, 1964; Grandcolas, 1994, 1996, 1999; Kambhampati, 1995, 1996; Klass, 1997, 1998; Nalepa and Bandi, 1999). Despite renewed interest in cockroach phylogenetics and a wealth of new data, major differences remain among proposed phylogenetic arrangements for cockroach families (Fig. 2). For divergences as old as those hypothesized for cockroach families, it is likely that the more slowly evolving regions of the nuclear ribosomal DNA (18S rDNA, Lo et al., 2000; 28S rDNA, this study) and conserved nuclear protein encoding genes could be important sources of new evidence on cockroach relationships.

Our sequence comparisons and molecular phylogenetic analyses are used to describe rDNA repeat unit variation at differing taxonomic levels. First, we evaluate the phylogenetic utility of a ~1200 bp fragment of the 28S rDNA for reconstructing higher-level cockroach relationships. Our findings show that the 28S rDNA is highly informative for higher-level cockroach phylogeny. Second, comparative analysis of the nucleotide sequences and the secondary structures of the ITS1 and ITS2 among closely related *Blattella* and *Periplaneta* species reveals major structural and sequence-level constraints.

2. Materials and methods

2.1. Cockroach taxa sampled

rDNA sequences were obtained from 11 cockroach species from three families: Blattellidae — *Blattella germanica*, *B. vaga*, *B. lituricollis*, *B. asahinai*, *Parcoblatta latta*; Blattidae — *Periplaneta americana*, *P. fuliginosa*, *P. brunea*; Blaberidae — *Diploptera punctata*, *Blaberus atropus*, *B. giganteus*. Specimens were obtained from laboratory cultures (CS lab, NCSU), or from colleagues, and frozen at −80 °C to preserve nucleic acids.

2.2. Laboratory methods

Total genomic nucleic acids were extracted using a standard DNA extraction buffer (Tris–HCl, proteinase K, SDS) according to the protocol described in Mukha et al. (1995).

The primers used for amplification and sequencing are shown in Fig. 1B. The priming sites for DAMS-18 and DAMS-28 are highly conserved in eukaryotes; these primers are “universal” and may be used for amplification of the corresponding rDNA fragments across a broad taxonomic range of organisms (Mukha and Sidorenko, 1995, 1996; Mukha et al., 2000). In addition, we developed seven additional priming sites, specific for cockroaches, within the rDNA repeat unit (coc1–coc7, Fig. 1B).

PCR amplification of overlapping fragments within analyzed rDNA fragments was carried out using Taq DNA Polymerase (Promega) and the PTC-100 Thermal Cycler (MJ Research Inc.). Each reaction contained 0.1 mg DNA template, 1.5 mM MgCl and 1 mM each dNTP. The PCR regimen was as follows: initial template denaturation at 95 °C for 5 min, followed by 30 cycles of 94 °C for 1 min, 55 °C for 2 min, and 72 °C for 3 min, and a final 7 min elongation cycle at 72 °C.

2.3. DNA sequencing

Sequences were obtained by dye terminator cycle sequencing using the ABI Taq FS enzyme (PE Applied Biosystems, Foster City CA), gel fractionated, and base-called on the ABI PRISM™ 377 DNA sequencer (PE Applied Biosystems, Foster City, CA) of the North Carolina State University DNA Sequencing Facility. Opposite strands were confirmed for all templates. ABI trace files were edited and contigs assembled using the program GAP4 in the STADEN software package (Staden, 1996) on a SUN Ultra1 workstation.

2.4. Nucleotide alignment and phylogenetic analysis

Alignments were constructed using the multiple sequence alignment algorithm in the program Clustal-W 1.7 (Thompson et al., 1994). The program FASTA (Pearson, 1990) was used to investigate pairwise sequence differences for ITS1 and ITS2 in closely related cockroach species. For the phylogenetic analysis of the 28S rDNA fragment, we analyzed two nucleotide alignment data sets. One included the entire optimal alignment generated in Clustal-W 1.7. For the second, we excluded positions within the 28S gene for which primary sequence homology was in doubt, thus treating

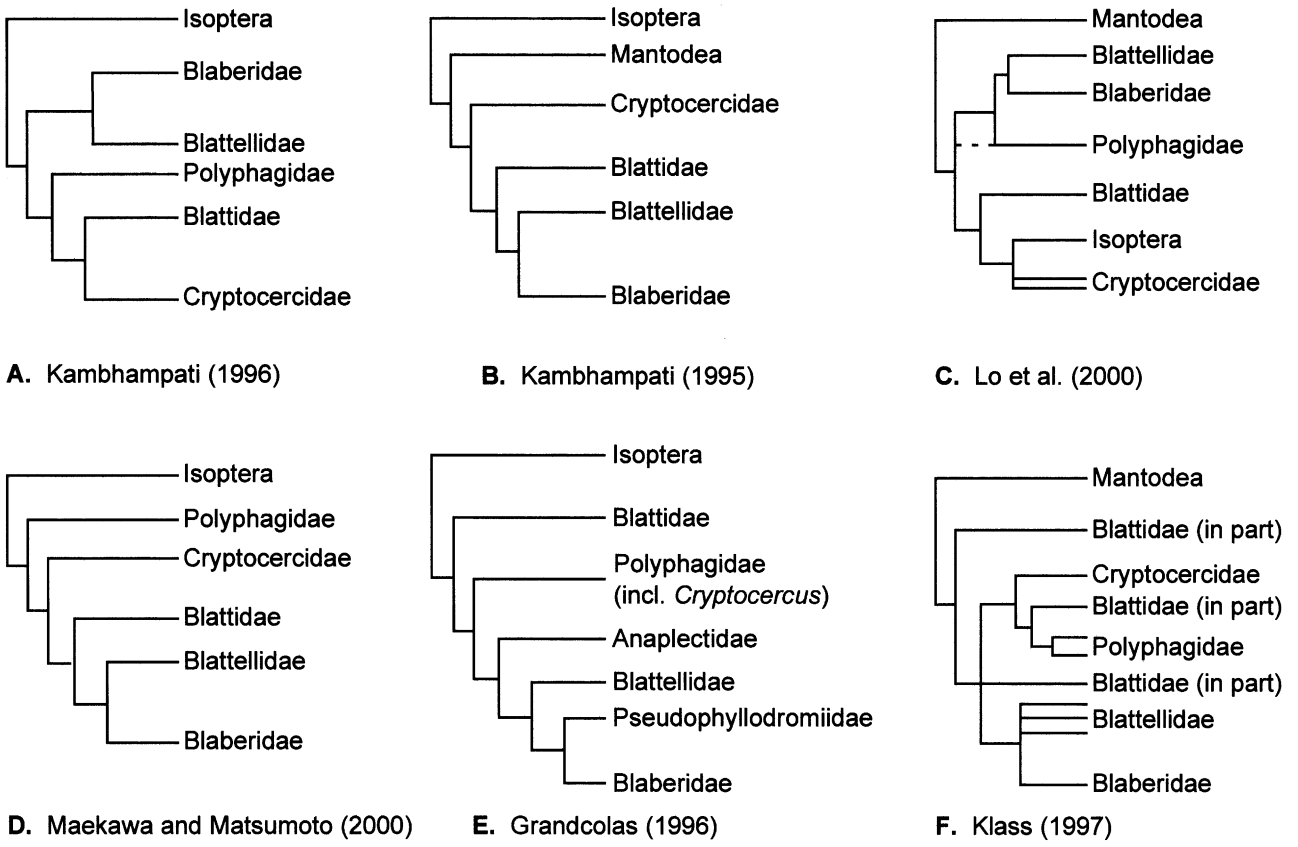


Fig. 2. Alternative phylogenetic hypotheses for the higher-level relationships of cockroaches. (a) 12S rRNA (Kambhampati, 1996); (b) 16S and 12S rRNA (Kambhampati, 1995); (c) multiple gene sequences (Lo et al., 2000); (d) mt COII (Maekawa and Matsumoto, 2000); (e) comparative morphology (Grandcolas, 1996); (f) comparative morphology (Klass, 1997).

as missing data positions for which alternative ad hoc placement of indels could affect the phylogenetic outcome. The latter method uses only those positions whose homology is certain to infer a tree. All of the alignments generated in the current study can be obtained from the following website: http://www2.ncsu.edu/unity/users/b/bwiegman/public_html/align.html and from the EMBL nucleotide database. Gaps in unambiguously alignable regions were also treated as missing data. Phylogenetic trees were constructed using parsimony and maximum likelihood methods in PAUP∗4b2a (Swofford, 1999). Parsimony searches were conducted using the branch-and-bound option in PAUP∗ (Swofford et al., 1996). Heuristic likelihood searches were conducted under the following model parameters: HKY85 (Hasegawa et al., 1985) model of nucleotide substitution with empirical base frequencies and transition/transversion ratio set to 2.0. Bootstrap support values (Felsenstein, 1985) were obtained from 1000 replicate re-sampled data sets for parsimony analysis and 500 replicates for maximum likelihood.

2.5. rDNA secondary structure estimation

We estimated secondary structures for the rRNA fragment containing 5.8S–ITS1 sequences. Folding patterns were compared in two closely related species: *B. germanica* and *B. lituricollis*. Zuker's dynamic programming algorithm (Zuker and Stiegler, 1981) was used to calculate secondary structures. Estimated structures were visualized using RNAdraw software (http://iubio.bio.indiana.edu/IUBio-Software+Data/molbio/ibmpc/rnadraw-readme.html).

3. Results and discussion

3.1. Phylogenetic utility of the 28S rDNA in Blattodea

Alignment of the 28S rDNA sequence data for the 11 cockroach species used in this study resulted in 883 sites included in the phylogenetic data set. Of these, 267 were variable and 156 were parsimony informative (see http://www2.ncsu.edu/unity/users/b/bwiegman/public_html/align.html). Analysis of the entire alignment generated in Clustal-W 1.7 with all positions included gave identical phylogenetic results to those reported below.

The length of the fragment was nearly identical for all of the sequenced taxa (Table 1). Uncorrected pairwise sequence divergence ranged from 0.1 to 8% for *Blattella* species, 11 to 19% between genera and families, with

Table 1
Length variation and nucleotide composition of sequenced cockroach rDNA regions

Species	rDNA Region Length (nt)			%GC			GenBank Acc number
	ITS1	ITS2	28S∗	ITS1	ITS2	28S[a]	
Blattella asahinai	667	443	873	52	55	38	AF321253
Blattella germanica	661	445	873	52	54	39	AF321244
Blattella lituricollis	841	493	873	54	55	38	AF321245
Blattella vaga	291	442	872	48	52	39	AF321246
Parcoblatta latta	393	711	874	53	61	36	AF321247
Blaberus atropus	342	282	871	58	61	36	AF321252
Blaberus giganteus	N/A	N/A	871	N/A	N/A	36	AF321254
Periplaneta Americana	391	267	874	52	65	35	AF321248
Periplaneta brunea	357	226	874	59	61	34	AF321249
Periplaneta fuliginosa	343	264	874	59	64	35	AF321250
Diploptera punctata	379	247	860	–	55	38	AF321251

[a] 28S lengths represent sequenced and aligned lengths of the 5′ fragment sampled in the current study only.

the most divergent comparisons between *Diploptera* and *Blattella*. Average base frequencies for this fragment were A=16.10, C=34.07, G=29.29, and T=20.06%, reflecting differential base substitution rates across deep divergence times. A Chi square test for base composition showed no significant deviation from these proportions among taxa (χ^2=19.26, d*f*=30, *p*=0.93). To assess whether the sequenced 28S rDNA fragment is saturated for the divergences among our taxa, total number of transitional and transversional changes were plotted against HKY model-corrected pairwise distances (Fig. 3). For both classes of substitution, numbers steadily accumulated as corrected pairwise divergence increased, indicating that saturation had not been reached (Fig. 3).

Parsimony analysis of the 28S rDNA data set yielded a single most parsimonious tree (Fig. 4; length=395; CI=0.85; RI=0.86). The resulting tree topology was well-

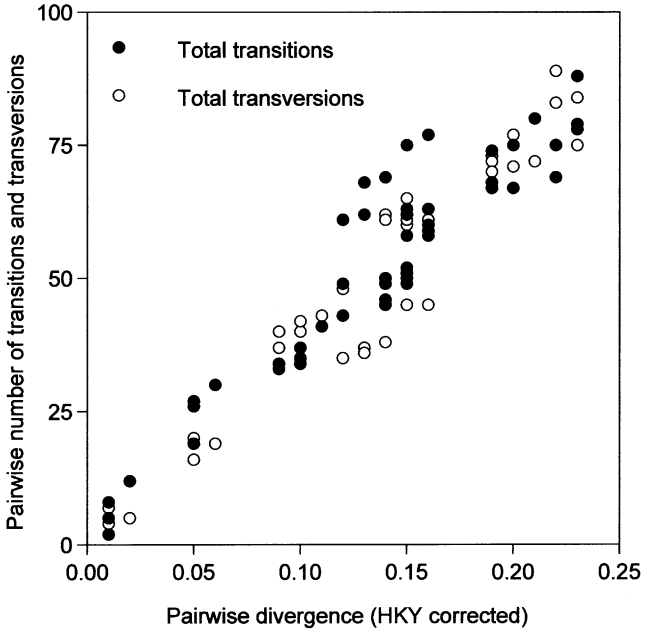


Fig. 3. Summed transitions and transversions in the 883 bp 28S rDNA fragment as a function of HKY-corrected sequence divergence for each pairwise comparison of taxa.

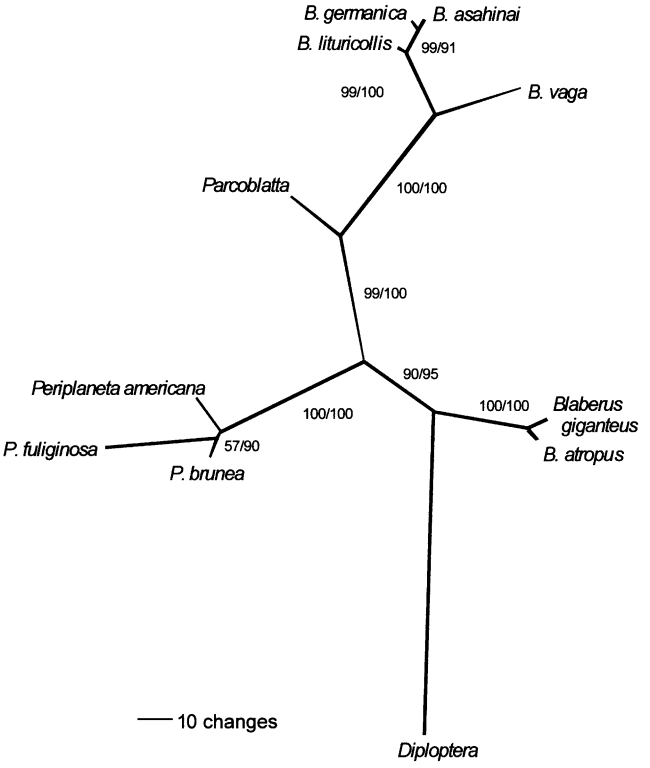


Fig. 4. Single most parsimonious tree and identical maximum likelihood topology inferred from the 883 bp 28S rDNA fragment (length=604; CI=0.92; RI=0.76; (−lnL=3250.52). Node support values above the line are bootstrap percentages based on 1000 replicate parsimony searches/500 replicate maximum likelihood searches. Branch lengths are proportional to the number of parsimony-based assigned character changes under ACCTRAN optimization in PAUP∗ 4b2a.

supported, with high bootstrap values (above 89%) supporting most of the major internal nodes. Maximum likelihood analysis of the same data yielded an identical tree with similar support values (−lnL=3250.52; Fig. 4). The relationships implied by this analysis are consistent with several published hypotheses of cockroach phylogeny (Kambhampati 1995; Grandcolas 1996; Fig. 2). Our tree supports the monophyly of the included *Blattella* species and a close relationship of these to *Parcoblatta* (Fig. 4). Support for previous findings based on mitochondrial RNA genes (Kambhampati, 1995) and morphology (Grandcolas, 1996, 1999), which placed Blaberidae in or near the Blattellidae, depends on identifying the root position for our tree. Placement of blaberid species (*Diploptera punctata*, *Blaberus atropus*, *B. giganteus*) within our unrooted network is consistent with previous molecular and morphological results, but cannot rule out the alternative placement as sister group to Blattidae implied by a root node position for the blattids between the Blattellidae and all remaining taxa. Sequences from additional taxa, notably Cryptocercidae and Polyphagidae, as well as from close relatives to cockroaches, Isoptera (termites), would provide useful information about the phylogenetic relationships of these primitive taxa.

Nevertheless, our small taxon sample was only intended to guage the utility of 28S rDNA as a potential source of phylogenetic information. Our finding of the same well-supported topology using both parsimony and maximum likelihood methods with strong support for nodes, as well as the observed intermediate levels of variation (11–19% pairwise sequence difference between families) among the taxa, demonstrates that this gene contains sufficient nucleotide variation to be highly informative of cockroach phylogeny. This finding supports similar conclusions for the nuclear rDNA genes in other orthopterpoid insects (Flook et al., 1999), and insect order- and family-level analyses (Whiting et al., 1997; Wiegmann et al., 2000). The 28S rDNA would be an excellent candidate gene for higher-level phylogenetic studies in Blattodea.

3.2. ITS interspecies variation

The results of multiple nucleotide alignments among *Blattella* species (*B. germanica*, *B. asahinai*, *B. lituricollis*, *B. vaga*) and *Periplaneta* species (*P. americana*, *P. fuliginosa*, *P. brunea*) are shown in Figs. 5 (ITS1 sequences) and 6 (ITS2 sequences). Multiple insertions and deletions, as well as numerous point substitutions, are revealed. The most extensive changes were found within ITS sequences of the *Blattella* species. The ITS1 sequences of *B. germanica–B. lituricollis* and *B. germanica–B. vaga* were found to be 67.2 and 30.0% identical, respectively. Significant length differences were also found between *B. germanica*, *B. lituricollis* and *B. vaga* ITS1 sequences — 661 b, 841 b and 291 b, respectively (Table 1; Fig. 5a). ITS2 sequences of the compared species were much more conserved. The ITS2 sequences of *B. germanica–B. lituricollis* and *B. germanica–B. vaga* were 79.8 and 52.8% identical. Sequence length was also more similar among *Blattella* ITS2 sequences — *B. germanica* (445 b), *B. lituricollis* (493 b) and *B. vaga* (442 b) (Fig. 6a). Only minor differences were found between the closely related sibling species (Ross, 1988) *B. germanica* and *B. asahinai* ITSs: 96.1 and 97.3% similarity within ITS1 and ITS2, respectively. The lengths of the *B. germanica* and *B. asahinai* ITS1 (661 b and 667 b) and ITS2 (445 b and 443 b) were also nearly identical (Table 1; Figs. 5A and 6A). In addition to primary sequence similarity these species also shared a translocation of the rDNA cluster from chromosome X to chromosome 12 in their common ancestor (Ross, 1988).

The general character of the evolution of the ITSs within *Periplaneta* species was similar to that described above for *Blattella*. The ITS1 sequences of *P. Americana–P. fuliginosa* and *P. americana–P. brunea* were 59.5 and 55.4% identical. ITS1 sequence length varied in these species as well: *P. americana*, *P. fuliginosa* and *P. brunea* — 391 b, 343 b and 357 b, respectively (Table 1; Fig. 5B). ITS2 sequences of the compared species were more conservative. The ITS2 sequences of *P. americana–P. fuliginosa* and *P. americana–P. brunea* were 71.1 and 64.0% identical, and their lengths for *P. americana*, *P. fuliginosa* and *P. brunea* ITS2 sequences were 267 b, 264 b and 226 b, respectively (Table 1; Fig. 6B). All of the sequenced taxa showed a slight excess of guanine and cytosine in the ITS regions, while the 28S gene showed an elevation in adenine and thymine (Table 1).

Comparisons of sequences among *Blattella* and *Periplaneta* species suggest that substantial differences have accumulated in the internal transcribed spacer regions in the time since these species diverged. ITS sequences have been applied in phylogenetic analyses at the population and species level in many organisms, including insects (Wesson et al., 1992; Schlotterer et al., 1994; Vogler and DeSalle, 1994; Muccio et al., 2000). Although our species sample is small for *Blattella* and *Periplaneta*, variation within these genera suggests that the ITS sequences should be useful for phylogenetic reconstruction at the species-level in cockroaches. These sequences are easily alignable across their full length, and exhibit potentially informative primary sequence and length variation (Figs. 5 and 6). Based on this finding, we are curently compiling data for a larger species-level phylogenetic analysis of *Blattella*.

Unlike ITS, the 5.8S subunit sequence is nearly invariant among all the cockroach species sampled. The short length (~140 bp) and high conservation of this sequence limit its utility for cockroach phylogenetics. The alignment with CLUSTAL-W 1.7 of the 5.8S rDNA

Fig. 5. Comparison of the nucleotide sequences in ITS1; similar nucleotides are indicated by asterisks. (A) *Blattella* species; (B) *Periplaneta* species. **1.** *B. germanica*; **2.** *B. asahinai*; **3.** *B. lituricollis* ; **4.** *B. vaga*; **5.** *P. americana*; **6.** *P. fuliginosa*; **7.** *P. brunea*. The nucleotides that affect rRNA folding are shown in bold.

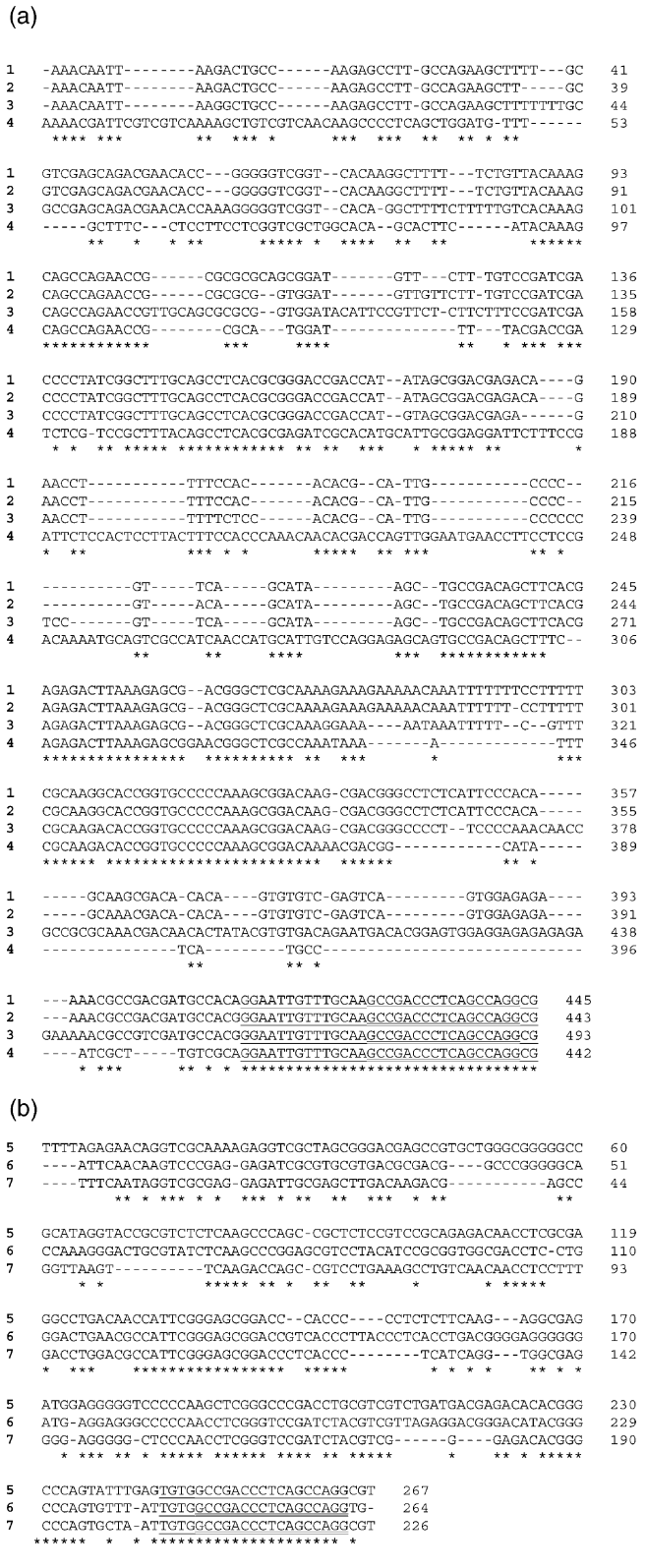


Fig. 6. Comparison of the nucleotide sequences in ITS2; similar nucleotides are indicated by asterisks. (A) *Blattella* species; (B) *Periplaneta* species. **1.** *B. germanica*; **2.** *B. asahinai*; **3.** *B. lituricollis*; **4.** *B. vaga*; **5.** *P. americana*; **6.** *P. fuliginosa*; **7.** *P. brunea*. Underlined and double-underlined sequences — evolutionarily conserved motifs (explanation in the text).

sequence data for the 11 cockroach taxa used in this study is available at http://www2.ncsu.edu/unity/users/b/bwiegman/public_html/align.html.

In addition to the observed variation, the ITS2 sequences also contained specific conserved motifs (underlined sequences in Fig. 6A and B, respectively). Elements of these motifs were shared between *Blattella* and *Periplaneta*, as indicated by double underlining in Fig. 6A and B. Additionally, a "super motif" — gccgaccctcagccagg — was shared by all of the Blattodea ITS2 sequences we have examined to date, and was also found in the published ITS2 sequences of more evolutionary distant insect species, such as *Nebria castanea* (Coleoptera), *Trichogramma dendrolimi*, *Leptopilina victoriae*, *Ganaspis xanthopoda*, *Melittobia digitata* (Hymenoptera), (GENBANK Accession numbers: AF173883, AF227949, AF015902, AF015892, MDU02950, respectively). It is commonly suggested that the high variability of rDNA spacer region sequences results from relaxed selection on the primary sequence, essentially that these regions are more "neutral" (sensu Kimura, 1983). However, evidence is accumulating that suggests that regions of sequence conservation in ITS sequences correspond to functionally important regions. The above described super motif is thus a likely candidate in the further identification of ITS structures that might play an important role in the function and organization of the insect rDNA.

3.3. ITS and secondary structure evolution

To infer whether ITS structural changes revealed here exert an effect on possible rDNA function, we compared secondary structures of the 5.8S rRNA compartment. This molecule is excised during processing and should fold as a separate unit, without the involvement of extended sequences of transcribed spacers (Gerbi, 1985). Folding of the 5.8S–ITS1 fragments of *B. germanica* rRNA is represented in Fig. 7A. The 5.8S rRNA did, in fact, fold as a separate compartment in this species. There was some minor variation in the observed size of the compartment among species; however, the inferred folding of the 5.8S rRNA when the ITS1 sequence was included corresponded to the fold inferred when 5.8S was analyzed without the ITS1 sequence (Fig. 7B). Folding of the *B. lituricollis* 5.8S–ITS1 sequence is represented in Fig. 7C. In *B. lituricollis*, sequences corresponding to 5.8S rRNA formed extended contacts with the ITS1 region, with an inferred folding pattern entirely different from that observed in *B. germanica*. In a hypothetical sequence from the 5.8S–ITS1 region of *B. lituricollis*, but with the sequences shown in bold in Fig. 5A removed (i.e., excluding a large insertion and a dinucleotide insertion present in comparisons of *B. germanica* and *B. lituricollis* ITS1 sequences), the basic *B. germanica*-like 5.8S rRNA fold was restored. Without these two

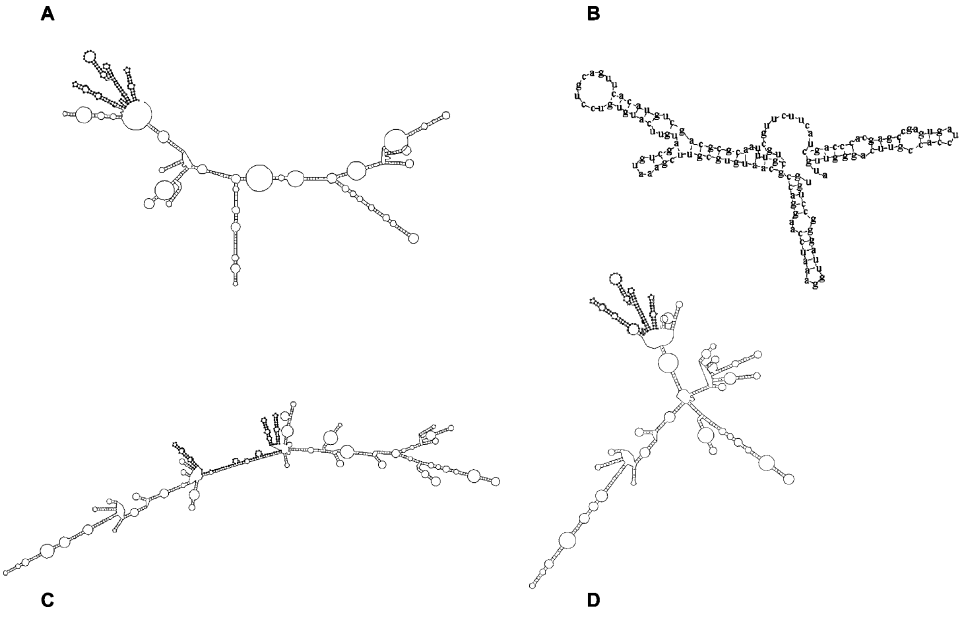


Fig. 7. Secondary structures of the rRNA regions based on DNA sequences. 5.8S–ITS1 fragment: (A) *B. germanica* 5.8S–ITS1 fragment; (B) *B. germanica* 5.8S fragment; (C) *B. lituricollis* 5.8S–ITS1 fragment; (D) *B. lituricollis* 5.8S–ITS1 fragment without nucleotides indicated in bold on Fig. 5A. The nucleotides corresponding to 5.8S rRNA were selected in A, C, and D.

insertions the 5.8S sequence now formed a discrete secondary structure domain separate from the ITS1 (Fig. 7D). These results suggest to us that the differences observed between ITS1 of *B. lituricollis* and its closest relative species are not "neutral" and are not simple accumulated random nucleotide changes, but bear a significant functional load. At least two multinucleotide insertions within ITS1 of *B. lituricollis* critically influence the folding of the rRNA precursor. Structural analysis of the ITS in yeast has shown that spacers play a role in the maturation of precursor rRNA molecules (Musters et al., 1990; Schulenburg et al., 1999; Van der Sande et al., 1992).

Critical changes in the rRNA folding pattern brought about by sequence evolution in the ITS spacer regions may thus have an important influence on the kinetics of precursor rRNA formation, and ultimately on the efficient functioning of the rDNA cluster.

Acknowledgements

We thank the North Carolina State University DNA Sequencing Facility for help with automated DNA sequencing. This study was supported in part by the Blanton J. Whitmire Endowment and the W. M. Keck Center for Behavioral Biology at NCSU, and by the Russian State Program "Frontiers in Genetics". We thank two anonymous reviewers for valuable suggestions.

References

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Flook, P.K., Klee, S., Rowell, C.H.F., 1999. Combined molecular and phylogenetic analysis of the Orthoptera (Arthropoda: Insecta) and implications for their higher systematics. Systematic Biology 48, 233–253.

Gerbi, S.A., 1985. Evolution of ribosomal DNA. In: McIntyre, R.J. (Ed.), Molecular Evolutionary Genetics. Plenum, New York, pp. 419–517.

Gerbi, S.A., Gourse, R.L., Clark, C.G., 1982. Conserved regions within ribosomal DNA: Location and some possible functions. In: Busch, H., Rothblum, L. (Eds.), The Cell Nucleus. Academic Press, New York, pp. 351–386.

Gonzales, I.L., Chambers, C., Gorski, J.L., Stambolian, D., Schmickel, R.D., Sylvester, J.E., 1990. Sequence and structure correlation of human ribosomal transcribed spacers. Journal of Molecular Biology 212, 27–35.

Grandcolas, P., 1994. Phylogenetic systematics of the subfamily Polyphaginae, with the assignment of *Cryptocercus* Scudder, 1862 to this taxon (Blattaria, Blaberoidea, Polyphaginae). Systematic Entomology 19, 145–158.

Grandcolas, P., 1996. The phylogeny of cockroach families: a cladistic appraisal of morpho–anatomical data. Canadian Journal of Zoology 74, 508–527.

Grandcolas, P., 1999. Systematics, endosymbiosis, and biogeography of Cryptocercus clevelandi and C. punctulatus (Blattaria: Polyphagidae) from North America: a phylogenetic perspective. Annals of the Entomological Society of America 92, 285–291.

Hamby, R.K., Zimmer, E.A., 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis, P., Soltis, D., Doyle, J. (Eds.), Molecular Systematics of Plants. Chapman & Hall, New York, pp. 50–91.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 21, 160–174.

Hillis, D.M., Davis, S.K., 1986. Evolution of ribosomal DNA: fifty million years of recorded history in the frog genus *Rana*. Evolution 40, 1275–1288.

Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. The Quarterly Review of Biology 66, 411–453.

Honda, J.Y., Nakashima, Y., Yanase, T., Kawarabata, T., Hirose, Y., 1998. Use of the internal transcribed spacer (ITS-1) region to infer *Orius* (Hemiptera: Anthocoridae) species phylogeny. Applied Entomology and Zoology 33, 567–571.

Kambhampati, S., 1995. A phylogeny of cockoaches and related insects based on DNA sequence of mitochondrial ribosomal RNA genes. Proceedings of the National Academy of Sciences USA 92, 2017–2020.

Kambhampati, S., 1996. Phylogenetic relationship among cockroach families inferred from 12S rRNA gene sequence. Systematic Entomology 21, 81–98.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, UK.

Klass, K.D., 1997. The external male genitalia and the phylogeny of Blattaria and Mantodea. Bonner Zoologische Beiträge 42, 1–341.

Klass, K.D., 1998. The ovipositor of Dictyoptera (Insecta): homology and ground plan of the main elements. Zoologischer Anzeiger 236, 69–101.

Kuperus, W.R., Chapco, W., 1994. Usefulness of internal transcribed spacer regions of ribosomal DNA in Melanopline (Orthoptera. Acrididae) systematics. Annals of the Entomological Society of America 87, 751–754.

Labandeira, C.C., 1994. A compendium of fossil insect families. Milwaukee Public Museum Contributions in Biology and Geology 88, 1–71.

Lo, N., Tokuda, G., Watanabe, H., Rose, H., Slaytor, M., Maekawa, K., Bandi, C., Noda, H., 2000. Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. Current Biology 10, 801–804.

Maekawa, K., Matsumoto, T., 2000. Molecular phylogeny of cockroaches (Blattaria) based on mitochondrial COII sequences. Systematic Entomology 25, 511–519.

McKittrick, F.A., 1964. Evolutionary studies of cockroaches. Cornell University Agricultural Experiment Station Memoirs 389, 1–197.

Michot, B., Bachellerie, J-P., Raynal, F., 1983. Structure of mouse rRNA precursors complete sequence and potential folding of the the spacer regions between 18S and 28S rRNA. Nucleic Acids Research 11, 3375–3391.

Mindell, D.P., Honeycutt, R.L., 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. Annual Review of Ecology and Systematics 21, 541–566.

Muccio, T., Marinucci, M., Frusteri, L., Maroli, M., Pesson, B., Gramiccia, M., 2000. Phylogenetic analysis of *Phlebotomus* species belonging to the subgenus Larrousius (Diptera: Psychodidae) by ITS2 rDNA sequences. Insect Biochemistry and Molecular Biology 30, 387–393.

Mukha, D.V., Sidorenko, A.P., 1995. Detection and analysis of *Tetrahymena pyriformis* 26S ribosomal DNA domain sequences, differing in degree of evolutionary conservation. Molecular Biology (Moscow) 29, 529–537.

Mukha, D.V., Sidorenko, A.P., 1996. Identification of high conservative domains within the 17S ribosomal DNA sequence from *Tetrahymena pyriformis*. Genetika 32, 1494–1497.

Mukha, D.V., Sidorenko, A.P., Lazebnaya, I.V., Zakharov, I.A., 1995. Structural variation of the ribosomal gene cluster within the Insecta class. Genetika 31, 1249–1253.

Mukha, D.V., Sidorenko, A.P., Lazebnaya, I.V., Wiegmann, B.M., Schal, C., 2000. Analysis of intraspecies polymorphism in the ribosomal DNA cluster of the cockroach *Blattella germanica*. Insect Molecular Biology 9, 217–222.

Musters, W.K., Boon, K., Van der Sande, C.A.F.M., Van Heerikhuizen, H., Planta, R.J., 1990. Functional analysis of transcribed spacers of yeast ribosomal DNA. European Molecular Biology Organization Journal 9, 3989–3996.

Nalepa, C.A., Bandi, C., 1999. Phylogenetic status, distribution, and biogeography of *Cryptocercus* (Dictyoptera: Cryptocercidae). Annals of the Entomological Society of America 92, 292–302.

Pearson, W.R., 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods in Enzymology 183, 63–98.

Raué, H.A., Klootwijk, J., Musters, W., 1988. Evolutionary conservation of structure and function of high molecular weight ribosomal RNA. Progress in Biophysics and Molecular Biology 51, 77–129.

Ross, M.H., 1988. Cytological studies of *Blattella germanica* and *Blattella asahinai*: I A possible genetic basis of interspecific divergence. Genome 30, 812–819.

Schlotterer, C., Hauser, M-T., Von Haeseler, A., Tautz, D., 1994. Comparative evolutionary analysis of rDNA ITS regions in *Drosophila*. Molecular Biology and Evolution 11, 513–522.

Schulenburg, J.H.G., Englisch, U., Wagele, J.W., 1999. Evolution of ITS1 rDNA in the Digenea (Platyhelminthes: Trematoda): 3′ end sequence conservation and its phylogenetic utility. Journal of Molecular Evolution 48, 2–12.

Staden, R., 1996. The Staden sequence analysis package. Molecular Biotechnology 5, 233–241.

Swofford, D.L., 1999. PAUP∗. Phylogenetic Analysis Using Parsimony (∗and OtherMethods). Version 4. Sinauer Associates, Sunderland, Massachusetts

Swofford, D.L., Olsen, G.J., Wadell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics, 2nd ed. Sinauer, Sunderland, Mass, pp. 407–514.

Tang, J., Toe, L., Back, C., Unnasch, T.R., 1996. Intra-specific heterogeneity of the rDNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Molecular Biology and Evolution 13, 244–252.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Van der Sande, C.A., Kwa, M., Van Nues, R.W., Van Heerikhuizen, H., Raue, H.A., Planta, R.J., 1992. Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. Journal of Molecular Biology 223, 899–910.

Vogler, A.P., DeSalle, R., 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. Molecular Biology and Evolution 11, 393–405.

Wesson, D.M., Porter, C.H., Collins, F.H., 1992. Sequence and secondary structure comparisons of ITS rDNA in mosquitoes (Diptera: Culicidae). Molecular Phylogenetics and Evolution 1, 253–269.

Wesson, D.M., McLain, D.K., Oliver, J.H., Piesman, J., Collins, F.H., 1993. Investigation of the validity of species status of *Ixodes dammini* (Acadri: Ixodidae) using rDNA. Proceedings of the National Academy of Sciences USA 90, 10221–10225.

Whiting, M.F., Carpenter, J.C., Wheeler, Q.D., Wheeler, W.C., 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Systematic Biology 46, 1–68.

Wiegmann, B.M., Mitter, C., Regier, J.C., Friedlander, T.P., Wagner, D.M., Nielsen, E.S., 2000. Nuclear genes resolve Mesozoic-aged divergences in the insect order Lepidoptera. Molecular Phylogenetics and Evolution 15, 242–259.

Zuker, M., Stiegler, P., 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Research 9, 133–148.